



# Color-based Fusion of MRI Modalities for Brain Tumor Segmentation

Nachwa Aboubakr, Mihaela Popova, James L. Crowley

## ► To cite this version:

Nachwa Aboubakr, Mihaela Popova, James L. Crowley. Color-based Fusion of MRI Modalities for Brain Tumor Segmentation. MICAD 2021 - 2nd International Conference on Medical Imaging and Computer-Aided Diagnosis, Mar 2021, Birmingham, United Kingdom. pp.89-97, 10.1007/978-981-16-3880-0\_10 . hal-03174069

**HAL Id: hal-03174069**

**<https://hal.science/hal-03174069>**

Submitted on 18 Mar 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Color-based Fusion of MRI Modalities for Brain Tumor Segmentation

Nachwa Aboubakr<sup>1</sup>, Mihaela Popova<sup>1</sup>, and James Crowley<sup>1</sup>

Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG,  
38000 Grenoble, France,  
`{first_name.last_name}@inria.fr`

**Abstract.** Most attempts to provide automatic techniques to detect and locate suspected tumors in Magnetic Resonance images (MRI) concentrate on a single MRI modality. Radiologists typically use multiple MRI modalities for such tasks. In this paper, we report on experiments for automatic detection and segmentation of tumors in which multiple MRI modalities are encoded using classical color encodings. We investigate the use of 2D convolutional networks using a classic U-Net architecture. Slice-by-slice MRI analysis for tumor detection is challenging because this task requires contextual information from 3D tissue structures. However, 3D convolutional networks are prohibitively expensive to train. To overcome this challenge, we extract a set of 2D images by projecting the 3D volume of MRI with maximum contrast. Multiple MRI modalities are then combined as independent colors to provide a color-encoded 2D image. We show experimentally that this led to better performance than slice-by-slice training while limiting the number of trainable parameters and the requirement for training data to a reasonable limit.<sup>1</sup>

**Keywords:** Tumor Segmentation, MRI, Modality Fusion, Medical Imaging

## 1 MRI Segmentation

Radiologists detect pathologies by visual inspection of X-rays, Computerized axial tomography (CAT) scans, and Magnetic Resonance Images (MRI). Unfortunately, competent diagnosis requires years of experience, and many common pathologies are misdiagnosed. MRI images, in particular, are difficult to interpret, as an accurate diagnosis can require adjustments to a number of parameters and the use of multiple MRI image modalities.

The automatic segmentation of MRI images offers a unique set of challenges. Pixels in each 2D image must be considered as part of a 3D volume as neighboring voxels provide contextual information that can be important for interpretation. This information can be lost when processing each slice independently. Approaches based on slice-by-slice segmentation of pathologies tend to ignore

---

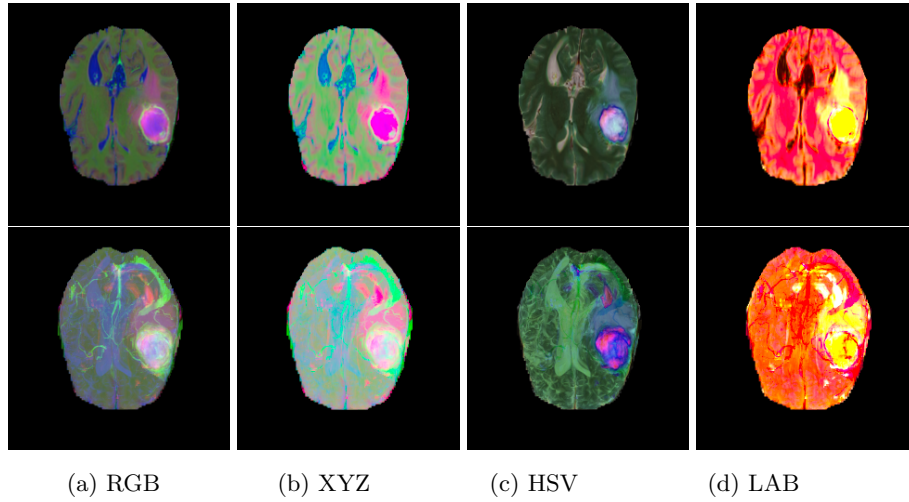
<sup>1</sup> This work has been partially supported by MIAI@Grenoble Alpes (ANR-19-P3IA-0003)

this information. On the other hand, 3D convolutional neural networks used for direct 3D segmentation require training large models as well as a considerable computational power and training data to converge. One possible approach to overcome this limitation is to transform the MRI 3D volume into 2D images using projections or slices at various angles [1, 2, 3]. However, different MRI modalities provide different information. Radiologists use multiple modalities when manually segment pathologies in MRI images.

Traditionally, radiologists use classical segmentation methods to segment different MRI modalities using techniques such as thresholding, region growing, edge detection, K-means [4, 5, 6]. With the rapid advances of artificial neural networks, a variety of MRI segmentation methods have been demonstrated to provide very promising tools to help radiologists. These approaches take an MRI image and produce a segmented image on areas of interest. Such approaches either process the whole 3D volume at once as in 3D U-Net [7] or they treat the MRI image slice-by-slice where each slice is processed independently. In some cases, slice-by-slice analysis is followed by intra-slice processing in a form of recurrent neural networks as in [8, 9, 10]. Approaches that treat the image slice-by-slice can be more affordable in terms of model size and number of parameters than approaches that treat the whole image at once. However, this generally comes at the cost of a reduction of performance.

Most work on automatic MRI segmentation either rely on a single MRI modality or consider each modality separately. To overcome the limitations of 2D approaches, several works have investigated ideas to transform the 3D MRI volume into 2D while keeping some kind of contextual information. These techniques include the use of multiple planes from different angles of an MRI image [3, 11, 12]. In addition, some recent works have investigated the use of MRI image projections from 3D to 2D using statistical measures such as Maximum Intensity Projection (MIP) [1, 2]. Few works have investigated the effect fusion of different medical imaging modalities. In [13], (CT, PET and MRI) images are color-fused which make them visually appealing and offer an accurate representation of the source images, and thus improving the diagnosis.

In this paper, we investigate the use of multiple MRI modalities for the automatic segmentation process, and examine the trade-off between computational cost and segmentation quality for different MRI modalities. To fuse MRI modalities, we encode each modality on a color channel and use encoded images as input for the automatic segmentation. We study the effect of using 2D projected images of the MRI volume instead of using the whole volume to minimize computational cost while preserving performance. We investigate an alternative approach for overcoming the limitations of 2D approaches by using Maximum Intensity projection of MRI volumes while exploiting information presence in different modalities by fusing them in color spaces.

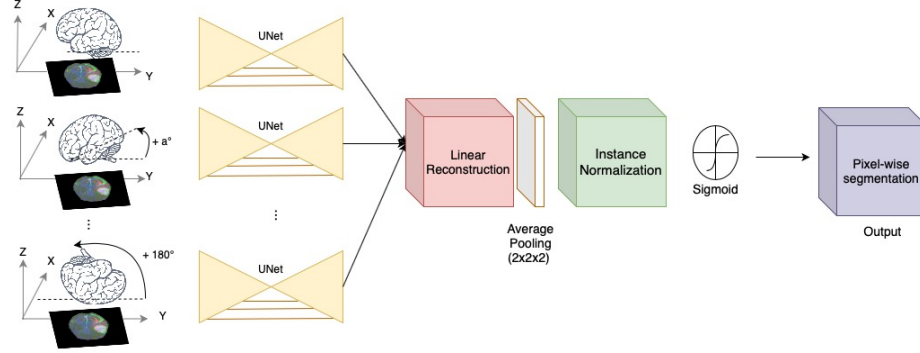


**Fig. 1.** Color-based fusion of MRI modalities. First row is slice-by-slice fusion. Second row is fusion of the Maximum Intensity Images of each modality. The fused modalities are (Flair, T1GD and T2) on different color-space channels, respectively.

## 2 Fusion of MRI Modalities

Different MRI modalities are employed for clinical diagnosis. These modalities include T1-weighted MRI (T1), T1 with contrast enhancement such as Gadolinium ions (T1GD), T2-weighted MRI (T2) and FLuid-Attenuated Inversion Recovery (FLAIR). Different MRI modalities show different information about the pathology. T1 shows healthy tissues with high intensity and the pathology with low intensity, T2 images represent pathology with high intensity. In T1GD images, the tumor border can be easily distinguished by the bright signal of the accumulated contrast agent in the active cell region of the tumor tissue. In FLAIR images, signal of water molecules is suppressed which helps in distinguishing edema region [14].

To exploit the knowledge that exists in different MRI modalities, we study the fusion of MRI modalities using a color space for brain tumor segmentation. We compare modality fusion using different color-space. These include RGB, XYZ, HSV and LAB. For the XYZ space; Y is the luminance, Z is quasi-equal to blue in RGB and X is the mix of three colors RGB, HSV is a color space that combines Hue, Saturation, and Values from different modalities and LAB is composed of L the luminance, A is a color value between green and red and B is a color value between blue and yellow. Of these spaces. RGB offers a linear fusion of the different modalities. The fusion with the other modalities is computed using color-space transformation formulas that transforms the fused image from the corresponding color space to RGB.



**Fig. 2.** Our model architecture. The input is a color-fused image of MRI modalities projected with Maximum Intensity projection. The 3D MRI image is rotated with "a" degrees. These projected images are input to a standard 2D U-Net. After that, the segmentation volume is reconstructed from the decoded U-Net output. After that, the segmentation volume is passed through refinement and normalization operations before producing the final pixel-wise segmentation.

*Model structure* , Our model architecture is shown in Figure 2. The base model is a standard U-Net model randomly initialized. The model receives fused 2D projected images extracted from MRI volumes of different modalities. The output of U-Net is used for 3D reconstruction of the images followed with a denoising layer. For training, we use a joint loss function of Dice loss and Cross-Entropy Loss, each contribute equally to the final loss. The implementation code of our method is available <sup>2</sup>.

*Preprocessing* , We extract multiple Maximum Intensity Images (MIP) from the 3D volume of MRI by rotating the volume from  $0^\circ$  to  $180^\circ$  around the axial plane with a steps and then project the resulting volume on the axial plane. We fuse MIPs from different modalities using a specific color-space. The resulted set of 2D images are used for training. We take color-fused MIPs of each angle and pass it to our model.

*Linear Reconstruction* , We use linear reconstruction to form a 3D tensor from 2D activation maps. Starting from an empty 3D tensor that matches the size of the 3D output mask, we add the first activation map to all the slices of the tensor. Then, we rotate the reconstruction tensor to match the angle of the new projection image, and we add the values of that projection to the resulted tensor from the previous computation. A voxel value in the result tensor is defined as the sum over the corresponding 2D projected values. This is repeated until all projections are added to the reconstruction tensor. Then the tensor is rotated one last time so that it goes to its original state.

<sup>2</sup> <https://github.com/Nachwa/Color-MRI-Seg>

Model	Dice	Precision	Recall	# Params
2D U-Net [17]	61.63	59.39	65.91	17M
Proj U-Net [1]	64.22	72.79	59.64	17M
2.5D U-Net [18]	64.98	62.86	67.26	-
3D U-Net [19]	<b>85.81</b>	<b>91.00</b>	<b>83.12</b>	51M
Ours	<u>79.20</u>	<u>79.06</u>	<u>80.94</u>	17M

**Table 1.** Segmentation results of our model with HSV color fusion compared to other competitive techniques. 3D U-Net and 2.5D U-Net results are reported in [19] and [18] respectively while the results of the other models are implemented.

*Denoising* , The linear reconstruction process produces artifacts and thus, the reconstructed volume needs denoising. The denoising process involves an average pooling layer with a kernel of (2x2x2). This is followed by an instance normalization layer and activated by a Sigmoid function to output the final predicted mask of the pathology.

### 3 Experimental Evaluation

For our experiments we used the BRATS 2017 dataset from the MICCAI Medical Imaging Decathlon [15, 16]. This dataset set provides Magnetic resonance images for the segmentation of brain tumor. The dataset contains four MRI modalities: Flair, T1, T1 with Gadolinium (T1GD), and T2. During our experiments, we use a fusion of Flair, T1GD and T2. The tumor/pathology has the highest intensity in T2 and Flair. In T1GD, the active part of the tumor has the highest intensity. We exclude T1 from our experiments since the pathology response to T1 modality has the lowest intensity which is incompatible with Maximum intensity projection.

The dataset provides labels for Edema, enhanced tumor, Non-enhanced tumor and Background. We consider for our experiments two classes only: The pathology which includes enhanced and non-enhanced tumor, and the background class which includes edema and background labels. For the evaluation of our experiments, we use four metrics: Dice score, Intersection over Union (IoU), Precision and Recall.

#### 3.1 Comparison with 2D/3D U-Nets

We compared our model with three already established models that can give us a good idea how well fusion and projections work in comparison with standard slice based or 3D image techniques. We compared with 3D U-Net with instance normalization and leaky ReLU which segments the whole 3D image [19]. Another method that we use for comparison is 2D U-Net [17] which segments each slice from the volume separately without taking into account the relationship between the slices. The third method that we used for comparison is 2.5D U-Net which

	Modality	Dice	IoU	Precision	Recall
Single	T1GD	72.50	60.56	77.78	72.08
	T2	63.80	51.81	70.59	63.68
	FLAIR	66.07	50.97	64.77	70.58
Fusion	RGB	76.46	65.17	<u>79.39</u>	78.34
	XYZ	76.00	62.92	<b>81.41</b>	73.36
	HSV	<b>79.20</b>	<b>67.04</b>	79.06	<b>80.94</b>
	LAB	<u>78.10</u>	<u>65.59</u>	77.51	<u>80.38</u>

**Table 2.** Tumor segmentation results of our model on different modalities. First rows are the results of each modality alone. Last rows are the results of fusing these modalities in corresponding color spaces.

takes a group of slices together for segmentation. When considering our method and these methods, we wanted to find a network that provided good quality of segmentation and does not need too many resources so that it can be used in real situation.

From Table 1, we found that 3D U-Net is performing the best for segmenting brain tumor in MRI. Our network placed second in terms of the quality of the segmentation after 3D U-Net. Our model outperforms, 2D U-Net which operates on slice by slice basis for the segmentation. Our model also outperforms both Proj U-Net and 2.5D U-Net which consider relations between slices either by using maximum intensity projections or using skip connections respectively. These results show that our model take advantage of the multi-modality fusion in addition of the intensity projection. Note that we report results of Proj U-Net using other modalities in Table 3 (top).

In terms of parameters 3D U-Net requires a high number of parameters to tune with 51 million. On the other hand, both our method and 2D U-Net require only 17 million parameters to tune. Following these experiments, we found that our proposed network provides a good balance between quality of segmentation and the training parameters needed.

### 3.2 Comparison with a single modality

In order to compare the differences in the quality of segmentation using a single modality and fusion of modalities, we compared the fusion of T1 with Gadolinium (T1GD), T2, Flair using RGB, XYZ, HSV and LAB color spaces. During these experiments, we have chosen against using T1 modality, because we use Maximum Intensity Projections and the pathology there has the lowest intensity while the healthy tissue has the highest intensity, so in that projection nothing significant can be observed in T1.

Comparing only the results for single modality, we can note that segmenting T1GD as input achieved the best results. This is likely the consequence of using the contrasting agent which highlights the active part of the tumor and could be

	Method	Dice	IoU	Precision	Recall
A	with T2	52.15	40.81	57.68	56.59
A	with T1GD	52.92	39.86	64.68	48.25
A	with Flair	64.22	49.73	72.79	59.64
B	A + RGB Color Fusion	72.23	58.32	74.93	72.20
C	B + Leaky ReLU	72.43	58.79	77.04	71.25
D	C + Instance norm	<b>76.46</b>	<b>65.17</b>	<b>79.39</b>	<b>78.34</b>
E	2D-UNet + RGB Fusion	67.09	55.69	64.83	70.35

**Table 3.** Ablation study shows the effect of our model structure choices. Method A (Proj U-Net) is implemented as described in [1]. Method D is our proposed model using RGB color fusion.

used in finding easily the boundary between the healthy tissue and the pathology which is a very important part in projection-based segmentation. On the other hand, the results for T2 and Flair alone are not as good, because on those images in addition to the pathology, the edema (swelling caused by the pathology) can also be seen with higher intensity than the healthy tissue and finding the boundary there would be more difficult.

From Table 2, we found that fusing the modalities in a color space makes a great difference in the quality of the segmentation. The use modalities fusion shows more details about different parts of the pathology. From our sample fusion images in Figure 1, we can see that the fusion in the HSV color space is visually more apparent than the other examples; in particular, we can notice that blood vessels in T1 with Gadolinium do not show in the image and the difference between the active tumor and the tumor core can be spotted easily.

From Table 2, we can see similar conclusion, fusion with HSV color-space outperforms the other color-spaces. RGB has an advantage over the other color spaces as it does not need any additional computation, so we can directly stack the projections of different modalities after normalization. For simplicity, we use RGB fusion for the rest of our experiments.

### 3.3 Ablation Study

In this section, we discuss our choice of the activation function and the normalization layer. In the related works [1, 2] they used ReLU as activation function and for normalization they choose Batch normalization.

When the activation function is ReLU, all negative values are reassigned to 0 while all positive values stay the same. The assigned of all negative values to 0 can lead to the vanishing gradient problem which can subsequently stop the network from training. One solution for this problem is to use leaky ReLU. For all negative values, leaky ReLU assigns them to the result of the multiplication of the value with 0.1 which removes the problem of neuron reaching 0 and dying. Thus, we wanted to experiment if changing the activation function would improve



the segmentation. The results from this experiment can be found in Table 3 (Row C) and we found that leaky ReLU improves the segmentation.

When working with medical dataset, we can expect that the pathology class would have less samples than the background, because the background contains the background of the medical image and the healthy tissue. In batch normalization, all images in the batch would be normalized together and, in our network, that would mean that all images from one patient are normalized together. Instance normalization normalizes each projection on its own. From Table 3 (Row D), we found that Instance normalization can significantly improve the segmentation of the pathology in all reported metrics.

We also compared the use of projected images against using the MRI all slices directly. Using maximum intensity projected images, the preprocessing time is 0.2 seconds for one patient, and additional half a second for the inference segmentation of the 3D image. On the other hand, segmentation using slice-by-slice MRI is 15 times more costly. Although, training slice by slice uses more data and takes more processing time, it does not improve the quality of the segmentation. From the results in Table 3 (Row E), we can notice that Modality fusion can improve the performance of standard slice-by-slice 2D-UNet with about 6 points. However, the use maximum intensity projection with the linear reconstruction improves the quality of the segmentation by about 10% over 2D-UNets.

## 4 Conclusion

In this paper, we investigate the fusion of Maximum intensity projected (MIP) images of MRI modalities using color spaces. We use MIP images of the MRI volume at different angles to minimize the processing time. We then color-fuse these projected on the RGB color space. In addition, we compare the performance of our model to 3D, 2.5D, and 2D U-Nets and show that our pipeline architecture provides a trade-off between performance and computational cost. We found that the use of modality fusion in a color space can improve the segmentation quality and the training time while preserving similar number of training parameters as 2D U-Net.

## References

- [1] Christoph Angermann, Markus Haltmeier, Ruth Steiger, Sergiy Pereverzyev Jr, and Elke Gizewski. Projection-based 2.5 d u-net architecture for fast volumetric segmentation. *arXiv preprint arXiv:1902.00347*, 2019.
- [2] Christoph Angermann and Markus Haltmeier. Random 2.5 d u-net for fully 3d segmentation. In *Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting*, pages 158–166. Springer, 2019.
- [3] Mathias Perslev, Erik Bjørnager Dam, Akshay Pai, and Christian Igel. One network to segment them all: A general, lightweight system for accurate 3d medical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 30–38. Springer, 2019.
- [4] Nick Eford. *Digital image processing: a practical introduction using java (with CD-ROM)*. Addison-Wesley Longman Publishing Co., Inc., 2000.
- [5] Nikhil R Pal and Sankar K Pal. A review on image segmentation techniques. *Pattern Recognition*, 26(9):1277 – 1294, 1993. ISSN 0031-3203. doi: [https://doi.org/10.1016/0031-3203\(93\)90135-J](https://doi.org/10.1016/0031-3203(93)90135-J). URL <http://www.sciencedirect.com/science/article/pii/003132039390135J>.
- [6] Song Yuheng and Yan Hao. Image segmentation algorithms overview. *arXiv preprint arXiv:1707.02051*, 2017.
- [7] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention*, pages 424–432. Springer, 2016.
- [8] Jianxu Chen, Lin Yang, Yizhe Zhang, Mark Alber, and Danny Z Chen. Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation. In *Advances in neural information processing systems*, pages 3036–3044, 2016.
- [9] Rudra PK Poudel, Pablo Lamata, and Giovanni Montana. Recurrent fully convolutional neural networks for multi-slice mri cardiac segmentation. In *Reconstruction, segmentation, and analysis of medical images*, pages 83–94. Springer, 2016.
- [10] Md Zahangir Alom, Mahmudul Hasan, Chris Yakopcic, Tarek M Taha, and Vijayan K Asari. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. *arXiv preprint arXiv:1802.06955*, 2018.
- [11] Sheng Hung Chung, Keng Hoon Gan, Anusha Achuthan, and Rajeswari Mandava. Liver tumor segmentation using triplanar convolutional neural network: A pilot study. In *10th International Conference on Robotics, Vision, Signal Processing and Power Applications*, pages 607–614. Springer, 2019.
- [12] Hasnine Haque, Masahiro Hashimoto, Nozomu Uetake, and Masahiro Jinzaki. Semantic segmentation of thigh muscle using 2.5 d deep learning network trained with limited datasets. *arXiv preprint arXiv:1911.09249*, 2019.
- [13] Karl G Baum, Maria Helguera, Joseph P Hornak, John P Kerekes, Ethan D Montag, Mehmet Z Unlu, David H Feiglin, and Andrzej Krol. Techniques for fusion of multimodal images: Application to breast imaging. In *2006 International Conference on Image Processing*, pages 2521–2524. IEEE, 2006.

- [14] Ali Işın, Cem Direkoğlu, and Melike Şah. Review of mri-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102:317–324, 2016.
- [15] Amber L Simpson, Michela Antonelli, Spyridon Bakas, Michel Bilello, Keyvan Farahani, Bram Van Ginneken, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv preprint arXiv:1902.09063*, 2019.
- [16] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4(1):1–13, 2017.
- [17] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [18] Jenny Stange Johansen and Mathias Aarseth Pedersen. Medical image segmentation: A general u-net architecture and novel capsule network approaches. Master’s thesis, NTNU, 2019.
- [19] Fabian Isensee, Jens Petersen, Simon AA Kohl, Paul F Jäger, and Klaus H Maier-Hein. nnu-net: Breaking the spell on successful medical image segmentation. *arXiv preprint arXiv:1904.08128*, 2019.